Niche Construction Affects the Variability and Strength of Natural Selection

Andrew D. Clark,[1,*] Dominik Deffner,[1,2,*] Kevin Laland,[1,†] John Odling-Smee,[3] and John Endler[4]

1. Centre for Biological Diversity, School of Biology, University of St Andrews, St Andrews, Fife KY16 9TH, United Kingdom; 2. Department of Human Behavior, Ecology, and Culture, Max Planck Institute for Evolutionary Anthropology, Leipzig 04103, Germany; 3. Mansfield College, Oxford University, Oxford OX1 3TF, United Kingdom; 4. School of Life and Environmental Sciences, Deakin University, Waurn Ponds, Victoria 3216, Australia

Submitted November 12, 2018; Accepted June 6, 2019; Electronically published December 6, 2019

Online enhancements: appendix, supplemental PDF. *Dryad data:* https://doi.org/10.5061/dryad.g66n3h5.

ABSTRACT: Consideration of the properties of the sources of selection potentially helps biologists account for variation in selection. Here we explore how the variability of natural selection is affected by organisms that regulate the experienced environment through their activities (whether by constructing components of their local environments, such as nests, burrows, or pupal cases, or by choosing suitable resources). Specifically, we test the predictions that organism-constructed sources of selection that buffer environmental variation will result in (i) reduced variation in selection gradients, including reduced variation between (*a*) years (temporal variation) and (*b*) locations (spatial variation), and (ii) weaker directional selection relative to nonconstructed sources. Using compiled data sets of 1,045 temporally replicated selection gradients, 257 spatially replicated selection gradients, and a pooled data set of 1,230 selection gradients, we find compelling evidence for reduced temporal variation and weaker selection in response to constructed compared to nonconstructed sources of selection and some evidence for reduced spatial variation in selection. These findings, which remained robust to alternative data sets, taxa, analytical methods, definitions of constructed/nonconstructed, and other tests of reliability, suggest that organism-manufactured or chosen components of environments may have qualitatively different properties from other environmental features.

Keywords: natural selection, selection gradients, niche construction, strength of selection, temporal, spatial.

Introduction

Field, experimental, and meta-analytical investigations convincingly demonstrate that quantitative trait variation is frequently under selection in natural populations (Endler 1986). The foundations for these advances were laid by Lande and Arnold (1983), who provided a rigorous theoretical framework for quantifying selection on phenotypic traits while controlling for selection on correlated traits, allowing for comparison of standardized coefficients across traits and study systems and for their use in meta-analyses to test general hypotheses (Endler 1986; Hoekstra et al. 2001; Kingsolver et al. 2001, 2012; Siepielski et al. 2009, 2013). Studies conducted over the past 40 years have generated many thousands of individual estimates of the direction and strength of natural selection in the wild, yet associated analyses of such environmental sources of selection and variation in selection are rare. Nonetheless, authors do commonly identify and informally discuss the environmental changes that are thought to have triggered the observed selective response. For example, significant variation in responses to selection is explained by precipitation (Siepielski et al. 2017), while more rapid phenotypic changes are reported in species exposed to human-disturbed (specifically, urban) compared to natural environments (Alberti et al. 2017). Variation in the magnitude and consistency of selection is partly attributable to differences in the focal taxa (e.g., plants vs. animals) or traits (e.g., life history vs. morphology), but there is still the general expectation that considering the properties of the agent or source of selection will help account for variation in responses to selection (Endler 1986; Linnen and Hoekstra 2009; Wadgymar et al. 2017).

Here we conduct meta-analyses of data from published studies of selection in the wild to test whether components of environments constructed by organisms differ from nonconstructed environmental components in the magnitude of and variance in selection generated (Laland et al. 2017). This expectation follows because organisms partly control their environments by building and regulating conditions in nests, burrows, mounds, webs, and so forth, which typically act to ensure that key environmental variables remain within suitable tolerance ranges (Odling-Smee et al. 2003; Hansell 2005, 2007; Sultan 2015). Organisms also choose habitats, resources,

* These authors contributed equally to this work.
† Corresponding author; email: knl1@st-andrews.ac.uk.

Am. Nat. 2020. Vol. 195, pp. 16–30.
DOI: 10.1086/706196

prey, foraging locations, nesting sites, oviposition sites, resting sites, and mates, and these decisions help ensure that the organism and/or its descendants experience suitable conditions (Huey et al. 2003; Odling-Smee et al. 2003; Sultan 2015; Stellatelli et al. 2018).

That organisms choose and modify environments in ways that have ecological and evolutionary consequences is uncontroversial (Jones et al. 1994, 1997; Odling-Smee et al. 2003, 2013; Armsworth and Roughgarden 2005; Edelaar et al. 2008; Ravigné et al. 2009; Bassar et al. 2010; Sultan 2015), and eco-evolutionary dynamics often focus on how organismal activities affect evolution by modifying fitness (Pelletier et al. 2009; Post and Palkovacs 2009; Bassar et al. 2010; Hendry 2016; see also earlier work on density-dependent selection, e.g., Pimentel 1961). However, there is still debate over whether the selective environments brought about by organisms are qualitatively or quantitatively different from selection arising from nonconstructed environmental sources (Scott-Phillips et al. 2014).

Here we test the prediction that environmental factors regulated by organisms (a.k.a. counteractive niche construction; Odling-Smee et al. 2003) will typically generate more consistent selection, sustained both over significant periods of time and across space (i.e., the same, or closely related, species should modify environmental components consistently over much of their geographical range), relative to nonconstructed environmental sources of selection (Laland et al. 2017). By nonconstructed sources we mean selection deriving from autonomous environmental processes that are not affected or are only weakly regulated by the activities of organisms (for discussion, see the appendix, available online). We test the prediction that where organisms regulate a source of selection, the observed response will be characterized by reduced temporal and spatial variance in selection gradients relative to nonconstructed sources. We further predict that the environment-buffering activities of organisms will reduce the strength of selection acting on a trait, leading to smaller magnitudes of linear selection gradients relative to nonregulated environmental sources.

These expectations follow from the assumption that the environmental regulatory traits of organisms have themselves been shaped by earlier selection (Laland et al. 2017). For instance, across diverse taxa, nests buffer climatic fluctuations that have been found to be metabolically costly, and nest building is thought to have evolved partly to fulfil an offspring insulation function (Cooper 1999; Hansell 2000; Reid et al. 2002; Odling-Smee et al. 2003; van Dijk et al. 2013; Combrink et al. 2017). Some species exploit the regulatory properties of other animals' constructions. Nile monitors, for instance, have evolved the behavior of laying their eggs inside termite mounds, thus exploiting the termites' thermoregulation of the nest (Hansell 1984). Irrespective of whether the inhabitants are the nest builder, their offspring, or inquilines, the conditions experienced by developing offspring are likely to be buffered relative to the external environment (Hansell 1984, 2000). Sometimes this regulatory activity is exhibited by multiple closely related species (e.g., *Platanthera* pollinated by several species of hawkmoth; Maad 2000). Similar logic

Table 1: Summary details of numbers of gradients with and without (in parentheses) borderline cases in our analysis of data sets comprising temporal variation in selection (pt. A), spatial variation in selection (pt. B), and spatiotemporal (i.e., pooled spatial and temporal) data (pt. C)

Category	Nonconstructed sources of selection	Constructed sources of selection	Mixed sources of selection	Combined (mixed + constructed) sources	Total
A. Temporal:					
Gradients	534 (488)	48 (48)	463 (443)	511 (491)	1,045 (979)
Subsets	94 (86)	16 (16)	58 (48)	74 (64)	168 (150)
Mean no. replicates	5.68 (5.67)	3 (3)	7.97 (9.23)	6.91 (7.67)	6.22 (6.53)
Studies	20 (18)	3 (3)	12 (11)	13 (12)	26 (23)
Species	19 (17)	3 (3)	11 (9)	12 (10)	24 (20)
B. Spatial:					
Gradients	102 (102)	71 (71)	84 (68)	155 (139)	257 (241)
Subsets	38 (38)	11 (11)	34 (26)	45 (37)	83 (75)
Mean no. replicates	2.68 (2.68)	6.45 (6.45)	3.00 (2.62)	3.44 (3.76)	3.10 (3.21)
Studies	9 (9)	2 (2)	3 (2)	5 (4)	11 (10)
Species	9 (9)	2 (2)	3 (2)	5 (4)	11 (10)
C. Spatiotemporal:					
Gradients	640 (594)	119 (119)	471 (451)	591 (571)	1,230 (1,164)
Studies	28 (26)	5 (5)	13 (12)	16 (15)	34 (31[a])
Species	26 (25)	5 (5)	12 (10)	15 (13)	31 (28)

[a] Two studies comprise both temporal and spatial data.

applies to macro- and microhabitat selection, which constitutes another means by which organisms can regulate the environmental conditions that they and their descendants experience (Huey et al. 2003; Sultan 2015; Stellatelli et al. 2018). When animals choose microhabitats with a suitable temperature, either for themselves or for their offspring, they too have acted to increase the likelihood that the temperatures experienced stay within a tolerable range and have potentially reduced temperature-related selection. While different in important respects (Edelaar and Bolnick 2019), environmental regulation and habitat choice are both activities that function to preserve an adaptive match between the features of organisms and factors in their environments, and these parallels lead some researchers to subsume both within the wider category of niche construction (Odling-Smee et al. 2003). Other researchers question the utility of such a broad conception, preferring to restrict niche construction to the construction of artifacts or the activities of a focal species (Scott-Phillips et al. 2014). By repeating our analyses for broad and narrow definitions of niche construction, we are able to explore whether habitat modification and choice generate equivalent effects on selection. Likewise, by separately analyzing cases in which relevant environmental regulation was undertaken by the focal species or another species and by a single or multiple species, we are able to evaluate whether and how these differences impact selection.

Methods

Data Sets

Data sets (summarized in table 1, with full data available in the Dryad Digital Repository; https://doi.org/10.5061/dryad.g66n3h5; Clark et al. 2019) were compiled using the following criteria.

Criterion i. We used Siepielski et al.'s (2009, 2013) data sets for our analyses, supplemented by additional studies that met Siepielski et al.'s criteria. For ecological validity, our analyses focused solely on studies that report standardized linear selection gradients from quantitative traits in wild populations (sensu Lande and Arnold 1983). We used selection gradients rather than selection differentials because selection differentials do not control for the influence of correlated traits and therefore potentially confound selection arising from a constructed source with selection arising from nonconstructed sources (and vice versa). Gradients are therefore likely to estimate effects of niche construction more reliably than differentials.

Criterion ii. We excluded cases of sexual selection. While mate choice is viewed as niche construction by some definitions, we omitted such cases, as (*a*) their inclusion would require disentangling natural (strictly, ecological) and sexual selection, as well as mate choice, from same-sex competition, which few studies do; (*b*) such cases are not easily categorized as environment buffering (or not); and (*c*) we did not want to conflate analysis of the evolutionary consequences of niche construction with any idiosyncratic properties of sexual selection.

Criterion iii. We included only studies on quantitative phenotypic traits showing continuous trait variation, avoiding cases of selection on categorical traits (sensu Siepielski et al. 2009, 2013), because the distributions are not directly comparable to those of continuous traits.

Criterion iv. We did not use gradients calculated using the breeder's equation, as it requires strong assumptions about the causation of fitness variation that are likely to be violated in studies of natural rather than artificial selection (Morrissey et al. 2010).

Criterion v. We also did not use gradients calculated using principal components because these represent a combination of traits and thereby potentially confound constructed and nonconstructed sources of selection.

Criterion vi. Following Siepielski et al. (2013), we used only studies that report standardized selection gradients with associated standard errors, which were necessary to compute the mean and variance of selection accurately while controlling for sampling error (Morrissey 2016).

Categorization Protocol

Selection gradients were categorized as responses to constructed versus nonconstructed sources of selection, utilizing and extending the implementation guidelines of Laland et al. (2017). All mentions of "constructed" or "nonconstructed" refer to the source of selection. Full details of the categorization procedure are given in the appendix, with justifications for each categorization decision given in the data set (available in the Dryad Digital Repository; https://doi.org/10.5061/dryad.g66n3h5; Clark et al. 2019). The principal points are summarized below.

Point i. We evaluated whether the source(s) of selection acting on the focal trait (e.g., chick tarsus length) was constructed or not (e.g., primarily affected by selection deriving from the nest or from other environmental factors) and make no predictions about the evolution of niche-constructing traits (e.g., nest building). Most authors of the studies included in our database explicitly stated what they assumed to be the primary source(s) of selection, which we accepted, assuming author expertise of the study system. If the source of selection was not reported, it was sometimes possible to infer the primary source(s) of selection, based on information reported in related or comparable studies (often by the same authors) and knowledge of the study system and trait function. If we were not able to identify the main source of selection with confidence, we did not include that trait in our analyses.

Point ii. Predictions were tested on a trait-by-trait basis, recognizing that, in a given study system, some traits could

be responses to constructed elements of the environment and others not.

Point iii. Identified sources of selection were categorized as constructed if they were dominated by environmental factors chosen or manufactured by organisms, such as nests, burrows, webs, pupal cases, soil conditions, or microenvironments, and regulated within bounds by organismal activity. Constructed sources were not restricted to the focal population (e.g., mites living in birds' nests experience the same damped temperatures as the birds). Conversely, environmental agents were categorized as nonconstructed in cases where the focal trait was not or was only weakly influenced by the niche-constructing activities and choices of organisms. Recognizing that environmental factors frequently comprise a complex of interacting elements, we restrict use of constructed to those contexts in which niche construction is thought to be the primary causal agent underlying selection and functions in a consistent regulatory manner (i.e., a single signal of orderly, directed, and sustained niche-constructing activities is present; see the appendix for details). Cases where the source of selection comprised multiple species were categorized as constructed only if the species collectively engaged in niche construction in a consistent and coherent manner. For instance, closely related species that engaged in similar forms of niche construction were treated as a single constructed source of selection, while, conversely, multiple species exhibiting different and mutually inconsistent activities, behavior, and/or preferences were categorized as nonconstructed sources. To account for definitional disagreements, we ran separate analyses in which the constructed environmental sources were produced by the focal species or another species and a single or multiple species. Some cases unambiguously comprised both constructed and nonconstructed elements, and these were categorized as mixed, with the expectation that the relevant measures would be intermediate. Analyses were conducted separately both with and without borderline cases, and ambiguous cases were not used. The presented results are from analyses excluding borderline cases, so as to take the most conservative approach.

Point iv. Odling-Smee et al. (2003) define niche construction broadly to include the consistent choices of animals for habitats (including flower sources among pollinators) and prey types, each of which may act as sources of selection. Likewise, Laland et al. (2017) state that their predictions should extend beyond the construction of physical artifacts and apply to animal choices. However, other researchers (e.g., Dawkins 2004) have argued for a narrower definition of niche construction that includes only physical artifacts or those modifications of the environment that are adaptations. We account for these disagreements by running analyses deploying both broad and narrow interpretations of constructed environmental sources, with the former but not the latter including cases of prey, resource, or microhabitat choice.

Point v. While our analyses typically compared gradients in constructed, nonconstructed, and mixed categories (henceforth, "separate category analysis"), due to the relatively low number of constructed cases, we also performed analyses in which we compared nonconstructed gradients with a combined category comprising all constructed and mixed gradients (i.e., with some constructed element; henceforth, "combined category analysis").

Point vi. Laland et al. (2017) make distinctive predictions for novel forms of environmental modification that generate de novo selection (a.k.a. inceptive niche construction) and those forms of niche construction that function to buffer existing environment variation (a.k.a. counteractive niche construction). However, only eight gradients were coded as novel in our database, which is insufficient to support analyses.

Temporal Variation. By temporal variation we mean interannual differences in selection on a given trait within a given population. We used the temporally replicated coefficients in the data sets of Siepielski et al. (2009, 2013) as the foundation for our data set, but we also considered ~300 additional studies. The latter were identified by reviewing the primary literature from March 2008 to June 2017 by searching the Web of Science database using the following search words deployed by Siepielski et al. (2009) and parts/combinations thereof: "spatiotemporal variation," "temporal variation," "annual variation," "multi-year selection," "long-term selection," and "fluctuating selection." See the appendix for details.

Spatial Variation. By spatial variation we mean differences between populations of the same species in different locations for a given trait. We used the spatially replicated linear selection gradients in Siepielski et al. (2013) as our data set. We again categorized the data into subsets: all selection gradients from the same study, species, and year, calculated with the same phenotypic trait and fitness measure, repeated across multiple populations. See the appendix for details.

Spatiotemporally Pooled Data Set. To explore the strength of selection more broadly, we analyzed the temporal and spatial data sets separately, in addition to a pooled data set that combined our spatial and temporal data. All presented results are mean absolute values of selection gradients as the expectation of the folded normal distribution defined by equation (5) (i.e., we ignore the signs of the gradients in this analysis).

Interobserver Agreement. To assess the reliability of the categorization procedure, two researchers (A. D. Clark and D. Deffner) independently coded all subsets of the data set. In each case, the coder was blind to the actual values of selection estimates. Both coders were neutral in that they had not previously published on the topic or invested in the issues.

Cohen's κ is a direct estimate of the between-rater agreement for categorical items between two independent raters and ranges from 0 to 1. It is more reliable than simple percent agreement calculations because it also accounts for agreements that are expected to occur by chance alone (Cohen 1960). The majority of cases could be reliably categorized as constructed, mixed, or nonconstructed using this protocol (Cohen's $\kappa = 0.91$ for temporal data sets and $\kappa = 0.95$ for spatial data sets). According to established benchmarks for the interpretation of κ, this level of agreement can be regarded as almost perfect (Landis and Koch 1977; Altman 1991). Where occasional disagreements arose, they almost exclusively concerned whether a case should be coded as constructed or mixed.

Data Analysis

Temporal/Spatial Variation. To compare variation in selection between constructed and nonconstructed categories while taking account of variation due to sampling error, we conducted mixed effects meta-analyses of the following form (Morrissey and Hadfield 2012; Morrissey 2016):

$$\beta_{i,j,k} = \mu_k + u_{j,k} + m_{i,j,k} + e_{i,j,k}, \tag{1}$$

where $\beta_{i,j,k}$ is the ith selection gradient estimate for the jth subset in the kth niche construction category (constructed or nonconstructed); μ_k is the average selection gradient of category k; $u_{j,k}$ is the deviation of the average selection gradient for subset j from the group average; $m_{i,j,k}$ represents the deviation between each selection gradient estimate and the true value due to sampling error; and $e_{i,j,k}$ is the residual, the deviation of the selection gradient at time i (temporal analyses) or space i (spatial analyses), from the mean of the jth subset. While $e_{i,j,k}$ is regarded as noise in most studies, here the term represents variation within subsets and thus quantifies temporal variation between different years or spatial variation between different localities in the same subset. Both the subset variance σ^2_u and the residual (i.e., temporal or spatial) variance σ^2_e are estimated as part of the model; the variance of sampling errors, in contrast, can be calculated from the standard errors (SEs) reported in the literature: $\sigma^2_{m_{i,j,k}} = \mathrm{SE}^2_{i,j,k}$. In addition to estimating different intercepts (means) for the different categories, we also allowed the subset variance σ^2_u and residual variance σ^2_e to differ between categories.

These models were fitted in a Bayesian framework with Markov chain Monte Carlo (MCMC) sampling using the R (ver. 3.4.0) package MCMCglmm (Hadfield 2010) with inverse-Wishart-distributed improper priors on both variances. We used 120,000 MCMC iterations, a burn-in period of 20,000 samples, and a thinning interval of 10. To aid interpretation of the results, we calculated two additional measures of temporal and spatial variation from the samples of the posterior distributions (Morrissey and Hadfield 2012; Siepielski et al. 2013; Morrissey 2016). First, we calculated the mean absolute differences between two randomly chosen years or locations within the same subset for each category. Morrissey (2016) specifies the mean absolute value of the difference between two independent draws from the same normal distribution as follows:

$$\mathrm{E}[|x_i - x_j|] = \frac{2}{\sqrt{\pi}}\sigma(x). \tag{2}$$

This measure quantifies how much selection is expected to differ between two randomly chosen years or locations in the same subset. Second, we calculated a consistency measure, as proposed by Morrissey and Hadfield (2012), separately for niche construction categories. This consistency (or repeatability) is the proportion of the total variation that is explained by variation between subsets:

$$\text{consistency} = \frac{\sigma^2_u}{\sigma^2_u + \sigma^2_e}. \tag{3}$$

The higher the consistency, the smaller the amount of temporal/spatial variation relative to other sources of variation that exist between different traits, populations, fitness measures, and so forth.

A More General Double-Hierarchical Model. By letting intercepts vary among subsets, the previous models take account of species differences (and all of the other variables subsumed under each subset) in the average strength of selection. Variances of selection gradients, however, were assumed to be equal among subsets, with the consequence that these analyses could not control for subset differences in variation of selection, which is our primary target of inference here. To overcome these limitations, we developed a more general multilevel framework that explicitly models subset-specific variances and thus allows us to also account for subset differences (i.e., differences between species, traits, etc.) in temporal/spatial variation. As before, we allowed varying intercepts for different subsets and let the model estimate the distribution of those intercepts with weakly regularizing (hyper)priors on both mean and standard deviation. On top of that, we now also let variances vary by subset and estimated their properties. Specifically, we assumed that subset-specific variances are distributed exponentially, with λ determining the rate of exponential decline (i.e., expected temporal/spatial variation $= 1/\lambda$). The supplemental PDF (available online) contains the full model specification and a more detailed description. We coded and sampled this model using Hamiltonian Monte Carlo in Stan (ver. 2.17.3; Carpenter et al. 2017; Stan Development Team 2018) with 120,000 iterations.

Estimating varying effects on scale parameters requires substantially larger sample sizes compared to location parameters, which is why we could run this model only for the temporal data set, and for not the spatial data set, and only for the combined analysis.

Strength of Selection. To investigate whether the mean strength of selection differs in response to constructed and nonconstructed sources of selection, we conducted random effects meta-analyses of the following form for each category (k) separately (Morrissey and Hadfield 2012):

$$\beta_{i,l,m,n,p} = \mu + u_{\text{species},l} + u_{\text{study},m} + u_{\text{fitness},n} + u_{\text{trait},p} + m_{i,l,m,n,p} + e_{i,l,m,n,p}, \quad (4)$$

where $\beta_{i,l,m,n,p}$ is the ith selection gradient estimate for species l, study m, fitness measure n, and trait p, respectively, which were included as random effects; m is the average selection gradient; $u_{\text{species},l}$, $u_{\text{study},m}$, $u_{\text{fitness},n}$, and $u_{\text{trait},p}$ are the deviations of the average selection gradient for species l, study m, fitness measure n, and trait p from the grand mean, respectively; $m_{i,l,m,n,p}$ represents the deviation due to sampling error; and $e_{i,l,m,n,p}$ is the residual. Models were again fitted in a Bayesian framework with the MCMCglmm R package, using the same default diffuse priors and the settings described above. Because we are primarily interested in the strength of selection irrespective of direction, we estimated mean absolute values as the expectation of the folded normal distribution, defined by the following equation (Morrissey 2016):

$$\bar{\mu}_{|x|} = \sqrt{\frac{2}{\pi}}\sigma(x)\exp\left(\frac{-\mu_x^2}{2\sigma^2(x)}\right) + \mu_x\left(1 - 2\phi\left(\frac{-\mu_x}{\sigma(x)}\right)\right). \quad (5)$$

In this case, ϕ is the mean of a normal distribution defined by μ_x and $\sigma^2(x)$ folded about the origin (Morrissey 2016).

For all analyses, visual inspection of the trace plots suggested good model convergence and no problematic autocorrelation, with convergence confirmed by Gelman-Rubin analysis (threshold: 1.1) unless otherwise specified (Gelman and Rubin 1992; Plummer et al. 2006; Ieno and Zuur 2015). The code necessary to reproduce all analyses is available on GitHub (https://github.com/andrewdclark/niche-construction-affects-natural-selection).

To evaluate whether the weaker effects observed in our spatial variation analyses compared to our temporal variation analyses reflected lower statistical power, we conducted 1,000 replicates in which we randomly downsampled the temporal data set to the equivalent size of the spatial data set (see table S8; tables A1–A3, S1–S8 are available online). To aid interpretation of our findings, we adopt the convention of referring to very strong (99% credible intervals for the two distributions do not overlap), strong (95% credible intervals do not overlap), and weak (90% credible intervals do not overlap) evidence.

Generality of Conclusions. We conducted six additional sets of analyses to evaluate the generality of our findings.

Generality check i. Initially, all analyses were performed with constructed and mixed cases combined in a single category called combined niche construction. We regard this as the most reliable analysis, as there were sometimes comparatively few gradients in the constructed category, whereas the number of gradients in the combined category was always adequate. We then repeated the analyses with separated constructed and mixed cases.

Generality check ii. Additional analyses evaluated the generality of the effects of organism-derived environmental regulation. These include deploying broad and narrow definitions of niche construction; the broad definition included both organism-mediated perturbations of the environment and the choices of animals (e.g., for prey, oviposition site, or flower choice among pollinators), while the narrow definition included as constructed only selective responses to physical artifacts, such as nests or burrows. To determine whether these differences affected the conclusions, we also separately analyzed cases in which relevant environmental regulation was undertaken by the focal species or by another species and a single or multiple species.

Generality check iii. To rule out the alternative explanation that studies measuring selection in response to constructed sources are different in some unknown respect from studies measuring selection in response to nonconstructed sources, we repeated the analyses using only the subset of studies that measured selection in response to both constructed and nonconstructed sources in the same species at the same time.

Generality check iv. To evaluate the sensitivity of our findings to coding error, we conducted 1,000 replicates in which we deliberately miscategorized 10% of the data, selected at random (see table S7).

Generality check v. Where the data set was biased in favor of birds, we repeated the analyses with combined categories (*a*) for birds alone and (*b*) for all other groups. However, because of either low Bayesian convergence or large credible intervals deriving from a reduced data set, it was not possible to conduct the analysis of birds alone reliably for the spatial data set.

Generality check vi. Given that a previous study (Kingsolver et al. 2012) found that phenological traits show weaker selection and that we note an excess of phenological traits in the nonconstructed category, where we analyzed temporal variation and strength of selection, we performed an additional analysis with phenological traits removed.

Results

Temporal Variation

Combined Niche Construction Categories Model. As predicted, residual variances, which represent temporal variance within subsets (i.e., within groups of selection gradients for

particular traits under the same conditions but in different years), differed substantially between categories, being much lower for combined cases than nonconstructed cases (table S1; fig. 1*A*). The expected difference between two randomly chosen years within the same subset was also smaller for combined cases than nonconstructed cases (table S1; fig. 1*B*). However, consistency was only slightly lower for nonconstructed cases than combined cases (table S1; fig. 1*C*), possibly because this measure is also affected by differences between conditions in between-subset variation (see "Discussion"). The analysis supports the expectation that the variation within subsets due to temporal fluctuations is substantially lower and reduced across years for traits affected by niche construction compared to nonconstructed cases.

Deploying the double-hierarchical temporal variation model, we find that most subsets show rather low temporal variation for both the combined (fig. 1*H*) and the nonconstructed (fig. 1*I*) categories but that, in line with our predictions, there are substantially more subsets with rather high variation among nonconstructed compared to combined constructed cases. Comparing expected temporal variation $(1/\lambda)$ across all subsets (fig. 1*G*) confirms that selection varies considerably less across different years in cases involving constructed sources of selection compared to only nonconstructed sources. The double-hierarchical model yields

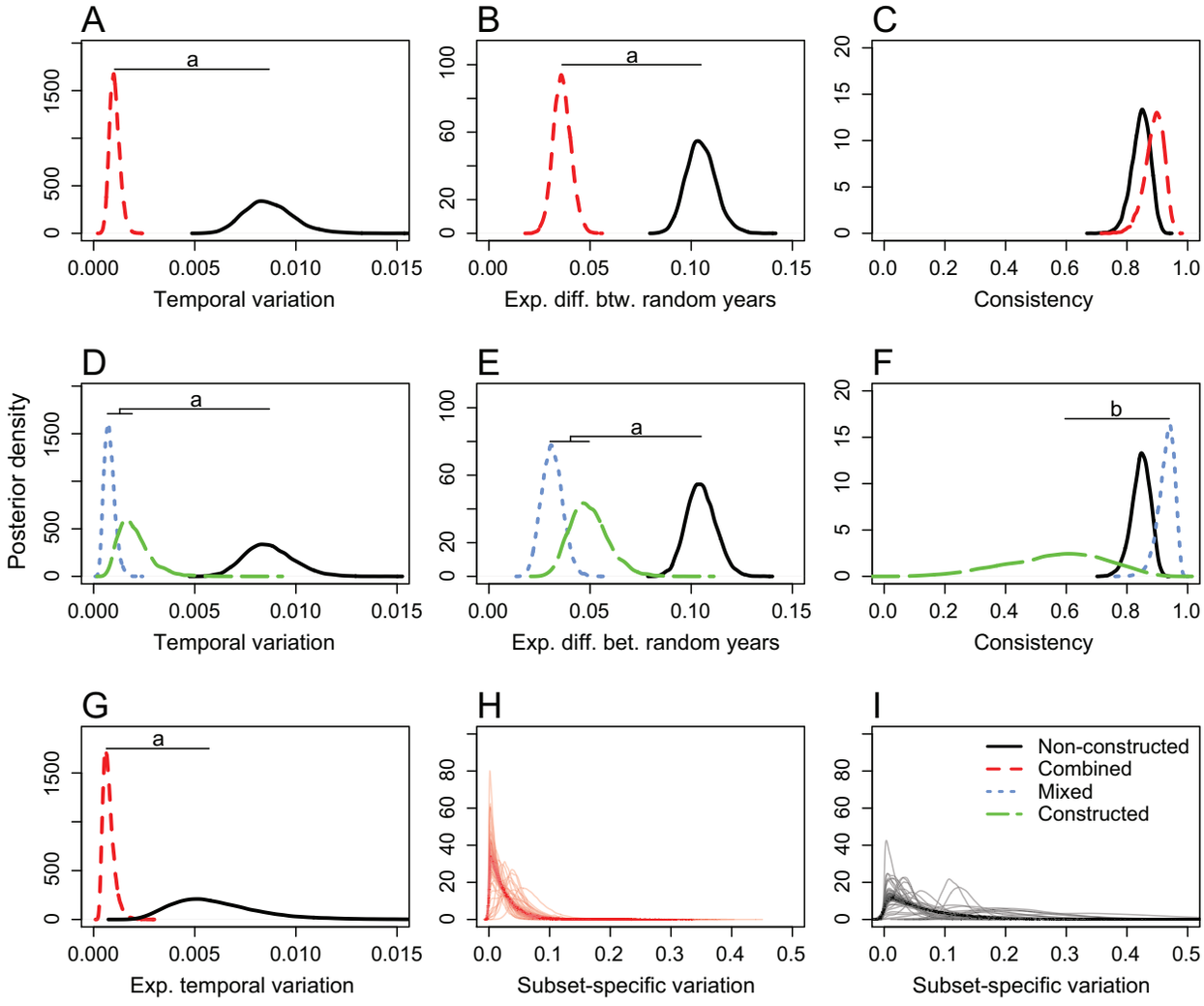

Figure 1: Marginal posterior probability distributions for temporal variation using the combined category (constructed + mixed) model (*A–C*) and the separate category model (*D–F*) and a broad definition of niche construction. *G–I*, Marginal posterior probability distributions for the double-hierarchical model. Expected temporal variation $(1/\lambda)$ was aggregated over all subsets (*G*) and standard deviations for each subset in the combined (*H*) and nonconstructed (*I*) categories. a = 99% credible intervals for the two distributions do not overlap; b = 95% credible intervals for the two distributions do not overlap.

the same qualitative results as the simpler model before and produces somewhat similar quantitative estimates, corroborating both the robustness of our previous results and the validity of the present model.

We conducted several checks of the generality of these findings (see table S2). (i) Repeating the analysis deploying a narrower definition of niche construction (physical artifacts only) generated similar results (fig. 2*A*). Whether the source of selection is the physical artifacts manufactured by organisms or derives from their environmental choices, temporal variation was smaller in the combined category than the nonconstructed category and was also more consistent and varied less across years. (ii) We also observed reduced temporal variation in the combined category compared to the nonconstructed category, irrespective of whether the focal species or another species engaged in the environmental regulation (fig. 2*B*). (iii) Likewise, irrespective of whether a single or multiple species engaged in the environmental regulation, similar patterns of reduced temporal variation were found in both cases (fig. 2*C*). (iv) Deliberately miscategorizing 10% of the data did not greatly affect the conclusions (with no overlap between the 95% credible intervals in 98.3% of simulations for within-subset variation and expected difference between years; see table S7). (v) Further analyses to check for taxon biases (figs. S2–S4; table S2; figs. A1, S1–S17 are available online) established that the general pattern of reduced temporal variation in selection from constructed sources does not arise from factors specific to birds or invertebrates. (vi) Restricting the analyses to studies that included selection estimates for both combined and nonconstructed categories again generated similar results (fig. S5; table S2), removing the possibility that the results could be attributed to a sampling bias in the choice of studies. (vii) Repeating the analysis without phenological traits (fig. S6; table S2) led to the same general trends but with relatively flat distributions for the nonconstructed cases and narrow distributions for the constructed cases that appear autocorrelated, raising questions about the reliability of these estimates in this reduced analysis. Analyses with and without borderline cases did not affect the findings.

Separate Niche Construction Categories Model. When partitioned into separate constructed and mixed categories, the

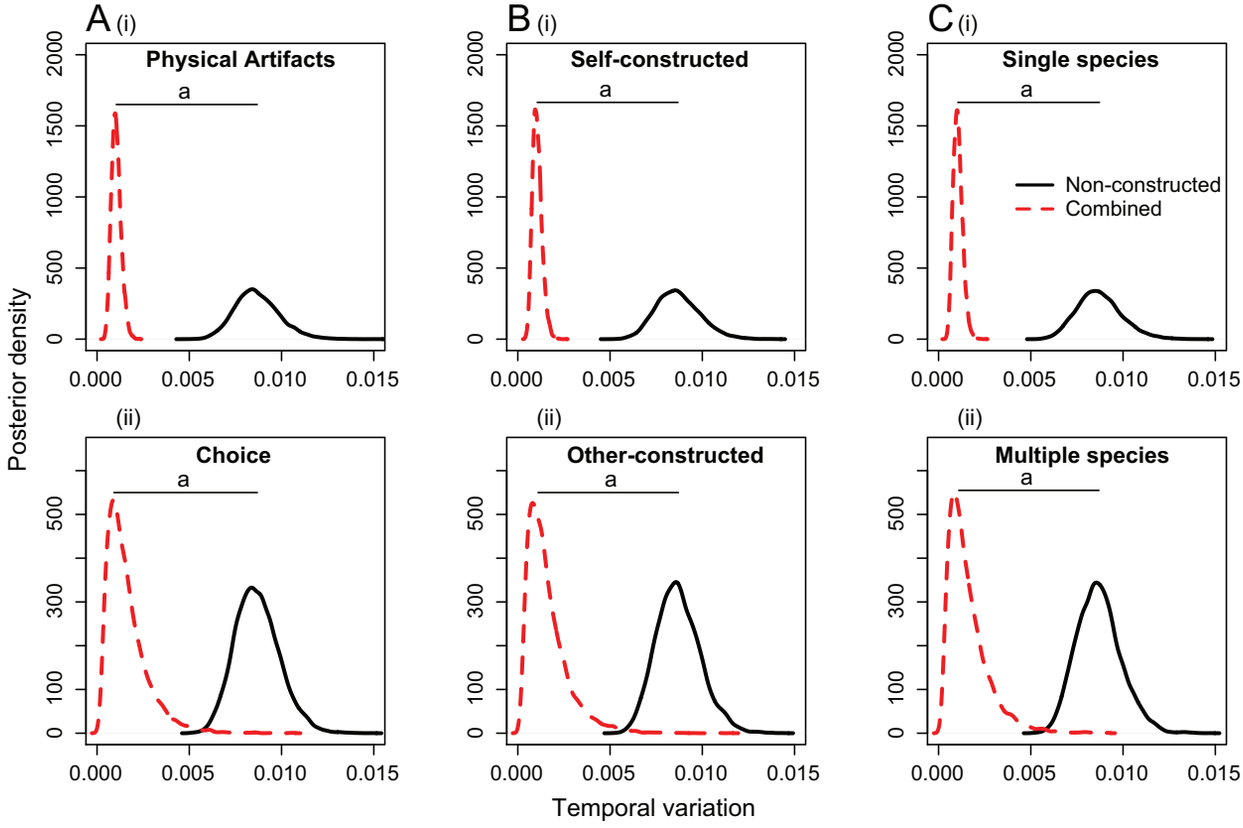

Figure 2: Marginal posterior probability distributions for temporal variation using the combined category model. Temporal variation is smaller in the combined (constructed + mixed) category than in the nonconstructed category, irrespective of whether niche construction (*A*) arises through the manufacture of physical artifacts or habitat choice, is by the focal species or another species (*B*), or involves a single or multiple species (*C*). a = 99% credible intervals for the two distributions do not overlap.

results largely confirm the above findings but with some interesting differences (table S1; fig. 1). As predicted, within-subset variation was greater in the nonconstructed category than in both the mixed and constructed categories (fig. 1*D*). The expected difference between two randomly chosen years within the same subset was also larger for nonconstructed cases than mixed and constructed cases, implying that niche construction is associated with a reduction in variability over different years (fig. 1*E*). Finally, while selection in response to mixed sources of selection was more consistent than selection in response to nonconstructed sources, the posterior distribution for constructed sources was very flat, indicating that the model could not reliably estimate consistency in this category, probably reflecting the lower sample size (fig. 1*F*). We again conducted generality checks of the findings, which largely confirmed earlier results but with some differences (table S3; figs. S7–S10). Taken together, our analyses deploying separate niche construction categories confirm that selection in response to constructed and mixed sources varies substantially less in time than selection in response to nonconstructed sources, although our expectation that the mixed category would sit between constructed and nonconstructed was not confirmed.

Spatial Variation

Combined Niche Construction Model. As predicted, residual variances, which represent spatial variance within subsets (i.e., within groups of selection gradients for particular traits under the same conditions but in different populations), also differed between categories, being higher for nonconstructed cases than combined cases (table S4; fig. 3*A*). The expected difference between two randomly chosen populations within the same subset was also smaller for combined cases than nonconstructed cases, but consistency was similar for nonconstructed and combined cases (table S4; fig. 3*B*, 3*C*). The analysis suggests that for traits affected by niche construction, the variation within subsets due to spatial differences is generally marginally lower, although not always more consistent across populations, compared to nonconstructed cases. However, it is hard to be completely confident in these results because of small sample sizes and large credible intervals.

We again conducted generality checks (table S5). Deploying the narrower definition of niche construction gave similar results (fig. S11), with spatial variation again smaller in the combined category than the nonconstructed category

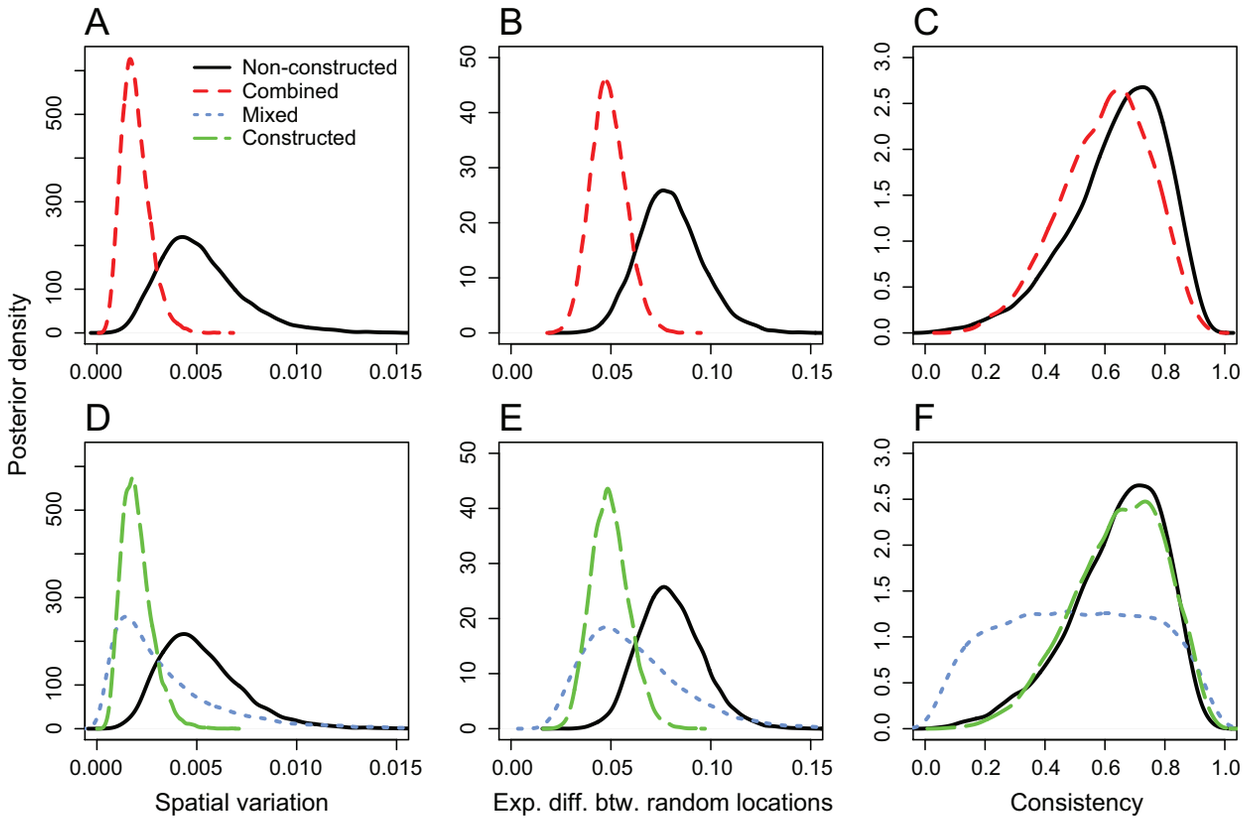

Figure 3: Marginal posterior probability distributions for spatial variation using the combined category model (*A*–*C*) and the separate category model (*D*–*F*) and a broad definition of niche construction.

but with the categories showing comparable consistency. Although repeating the analysis focusing solely on birds was not possible because of the small sample size (only four nonconstructed cases and with no cases categorized as mixed among nonbirds), analyses focusing on all taxonomic groups excluding birds produced the same results (fig. S12). This implies that the reduced spatial variation of selective responses to niche construction does not arise from factors specific to birds but holds equally in other groups of animals and plants. Contrary to our expectations, the residual spatial variance within subsets and expected difference between years were greater and consistency was lower in combined cases than nonconstructed cases when we restricted the analyses to all studies that included selection estimates for both categories (fig. S13). However, we are wary of placing much emphasis on these findings given the small sample size and often large credible intervals, as few spatially variable studies had gradients classified in both the combined and nonconstructed categories. This inference is supported by analyses in which we downsampled our temporal data set to the size of our spatial data set, which led to differences between categories being lost in 28% of cases (see table S8). Analyses with and without borderline cases did not affect the findings.

Separate Niche Construction Categories Model. As predicted, estimates of variation within subsets were higher in the nonconstructed category than in both the mixed and constructed categories (table S4; fig. 3*D*). The expected difference between randomly chosen sites within the same subset was larger for nonconstructed cases than mixed and constructed cases (table S4; fig. 3*E*), implying that niche construction is associated with a reduction in variability over different populations. While consistency was similar in nonconstructed and constructed sources, it was lower for mixed sources of selection and had a very large credible interval (table S4; fig. 3*F*). We note that the between-subset variance was lower for mixed cases and constructed cases (table S4) than nonconstructed cases, implying that selection in response to constructed sources of selection was remarkably similar across different traits and studies. Unfortunately, the small number of gradients precluded generality checks for the separate categories of analysis of spatial variation.

Strength of Selection

As predicted, the combined category had a lower mean absolute selection gradient value than the nonconstructed category (table S6; fig. 4*A*) but with partly overlapping distributions. When the analysis was repeated on the separate categories, the constructed category had a substantially lower mean than the nonconstructed category, with the mixed category intermediary, again as expected (table S6; fig. 4*B*). To check for a potential bias caused by the excess of phenological traits in the nonconstructed category, we repeated these analyses with phenological traits removed. In this reduced analysis, the combined category again had a lower mean absolute selection gradient value than the nonconstructed category (table S6; fig. 4*C*), with the overlap in distributions reduced relative to the above. When the analysis was repeated on the separate categories, constructed had a lower mean than nonconstructed, with mixed intermediate and the overlap in distributions again slightly reduced (table S6; fig. 4*D*). This implies that the phenological data may bias the findings, reducing the differences between category means. Splitting the data into separate temporal and spatial data sets produced almost identical findings in both cases, although the distribution was almost flat for the constructed category in the spatial separate analysis. Analyses with and without borderline cases did not affect the findings.

Discussion

While measures of natural selection in the wild are extensive, surprisingly little is known about the ecological factors underlying this selection (Endler 1986; Wade and Kalisz 1990; Caruso et al. 2017). Focusing solely on those organismal activities that buffer environmental variation, our study tested the predictions that natural (ecological) selection deriving from organism-constructed sources will exhibit reduced temporal and spatial variation in selection gradients and weaker (i.e., reduced intensity of) selection compared to nonconstructed sources (Laland et al. 2017). Our results support these predictions, with conclusions (i) broadly consistent across combined and separate niche construction category analyses and holding (ii) for both broad (including animal choices) and narrow (physical artifacts only) definitions of niche construction, (iii) whether the focal species or another species engages in the niche construction, (iv) for cases in which a single or multiple species engage in functionally equivalent niche construction, (v) in analyses that included all studies with selection estimates for both categories (undermining any suggestion that the findings reflect bias in choice of studies), and (vi) little affected by taxonomy. Evidence for weaker selection arising from constructed sources was perhaps more compelling in the separate niche construction category analysis than combined category analysis, with partial overlap in distributions. Further investigation provided some evidence that differences between categories in the numbers of phenological traits may have reduced the differences between categories, although only marginally.

These findings are in keeping with how natural selection is expected to shape environment-modifying and environment-choosing adaptations (i.e., niche construction) in ways that allow organisms to control and regulate features of their environment (Lewontin 1983, 2000; Odling-Smee et al. 1996, 2003; Sultan 2015). Such environment-buffering activities

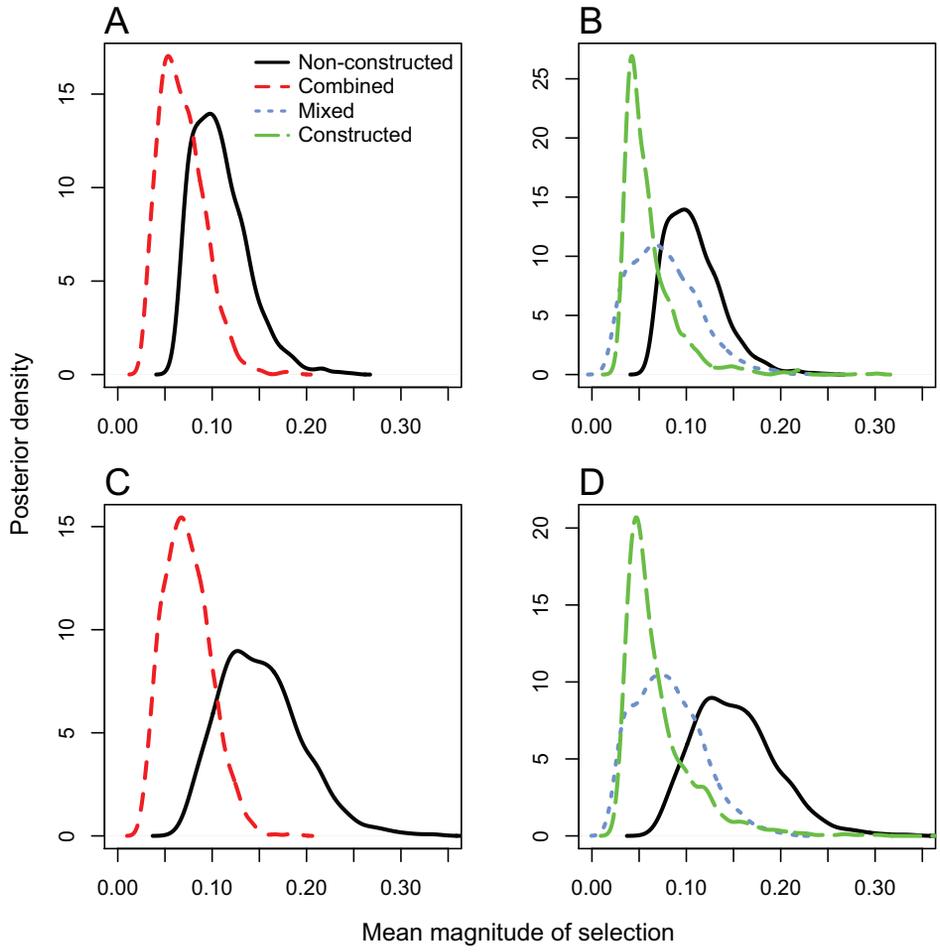

Figure 4: Marginal posterior probability distributions for the mean magnitude of selection using the pooled data set for the combined analysis (*A*), the separate analysis (*B*), the combined analysis without phenological traits (*C*), and the separate analysis without phenological traits (*D*).

help ensure that relevant resources are available to the constructors and their descendants and/or that they experience conditions held within suitable bounds. The fact that regulation through physical artifacts and habitat selection, by a single or multiple species or by the focal species or another species, generates similar impacts on selection supports the hypothesis that the functional equivalence of regulatory activity is often more important than the precise identity of the constructor in determining the selection that ensues and helps justify a broad definition of niche construction (Odling-Smee et al. 2003). The parallel effects on selection observed here should not be taken to imply that variant forms of niche construction will always have identical ecological and evolutionary causes and consequences. For instance, we anticipate that habitat modification will typically, although not inevitably, generate more profound ecological and evolutionary consequences for other species than habitat choice. Nonetheless, it is precisely because the parallels between these processes are not always self-apparent that the more general term of "niche construction" can have explanatory utility and brevity.

We note marginally stronger differences between conditions when the source of selection is constructed by the focal species rather than another species (fig. 2), which makes sense, as the focal species' niche construction has likely evolved to regulate environmental conditions for itself, while the regulatory activities of other species are tuned to enhance the other species' fitness. Plausibly, organisms will regulate the environment to reduce selection only, or more effectively, when they are already well adapted. However, since we examine evolutionary responses to niche construction rather than the evolution of the regulatory activity, loss of genetic variation in the regulatory activity does not account for the weaker selection.

While regulatory niche construction was predicted to lead to weaker and substantially lower variability in response to selection deriving from constructed sources compared to selection

arising from nonconstructed aspects of the environment (Laland et al. 2017), and while our findings uphold these predictions, there are other circumstances under which Laland et al. (2017) expect niche construction to increase the magnitude of selection. These include instances where organisms have recently modified environments in novel ways (inceptive niche construction; Odling-Smee et al. 2003). We were unable to find sufficient examples in the literature to conduct analyses to test this expectation, but, clearly, analyses showing a bidirectional pattern of response that is consistent with predictions would strengthen confidence in the conclusions.

Our study provides evidence that organism-driven modifications of the environment differ in important respects to other (nonconstructed) changes in the environment. While these findings will be intuitive to many researchers, they have been disputed (Scott-Phillips et al. 2014), and so an empirical demonstration was required. In fact, the idea that regulatory behavior might constrain evolution has been sporadically referred to in a literature that dates back at least to Bogert (1949) and his suggestion that reptiles can buffer selection through their behavior. More recently, other studies have explored this Bogert effect and report findings in keeping with our results (Huey et al. 2003; Stellatelli et al. 2018). While not testing this possibility directly, Siepielski et al. (2017) note that species may dampen the impact of climate-induced selection through migration and range shifts. Our study implies that phenomena akin to the Bogert effect may be more general than previously appreciated. Organism-derived regulation that buffers selection extends beyond direct behavioral adjustments to encompass buffering mediated by artifacts and other environmental resources and beyond animals to potentially any species that experiences environmental resources or conditions regulated by one or more organisms. Whether the consequences of regulatory activity are restricted to selection on the focal trait or exert a more general effect through impacting variance in fitness remains to be established. However, consideration of humans, who exhibit both extraordinary capabilities for environmental regulation and who have been subject to strong recent selection (Hawks et al. 2007), suggests that niche construction more likely does not buffer selection completely but rather shifts selection to and potentially intensifies selection on other traits—intuitions anticipated by Lewontin (1983) and Odling-Smee et al. (2003). Our findings demonstrate that niche construction impacts selection in the wild, but they do not prove that it imposes a statistical bias on evolution (Odling-Smee et al. 2003; Laland et al. 2017); nonetheless, a failure to find differences in selective responses to constructed versus nonconstructed sources would have undermined that hypothesis.

We report strong evidence for reduced temporal variation in response to constructed environmental resources compared to nonconstructed sources. While a broadly similar pattern is observed for spatial variation, we nonetheless view the spatial results as less reliable for two reasons. First, there was a very small sample of spatially variable selection gradients that met our criteria for reliable categorization as constructed or nonconstructed. Analyses downsampling the temporal data set support the conclusion that our spatial analyses had substantially lower statistical power than both our temporal analyses and previous analyses of spatial variation in selection (e.g., Siepielski et al. 2013). Second, Siepielski et al. (2013) point out that spatial variation in selection gradients is most likely nonrandom (and probably overdispersed relative to random expectations), and hence our comparison of constructed and nonconstructed cases probably does not tap into the true spatial pattern for natural systems. This may partly mask the signal of expected differences between construction categories if divergent spatial conditions imposed on natural systems expose organisms to conditions beyond the bounds of what they can buffer. These factors leave our spatial analysis less reliable than the temporal analysis, and we place less emphasis on it.

When looking at the differences in strength of selection between categories, we find modest but consistent differences between the strength of selection in constructed and nonconstructed cases. An analysis comparing coefficient standard deviations with those for absolute values (see figs. S16, S17) confirmed that the finding of less variation in selection associated with constructed compared to nonconstructed sources relates primarily to strength differences rather than to differences in the direction of selection. As expected, pure constructed cases experience a low strength of selection, but when autonomous elements also come into play in the mixed category, selection experienced by the traits becomes stronger. Here we note a possible bias generated by our coding system. Mixed cases were categorized as mixed only if they had a clear autonomous influence in addition to a constructed influence, which may have led to the neglect of mixed cases with weaker autonomous selection components and artificially inflated estimates of selection magnitude. In addition, combined and nonconstructed cases exhibited a substantial difference in the distribution of trait types, with a higher proportion of phenology traits (e.g., timing of life cycle events) compared to other traits in nonconstructed cases. This is potentially significant, as Kingsolver et al. (2012) found relatively weak selection on phenology compared to morphological traits. When phenological traits were removed from the data set, we indeed found that the mean strength of selection for the different categories did separate more in the predicted directions. Given that we had an excess of phenological traits associated with the nonconstructed category, this implies that the reduced selection on phenological data may have partly masked differences between niche construction categories.

One unanticipated difference between constructed and mixed cases lay in the smaller between-subset variance in constructed cases compared to both mixed and nonconstructed

cases, which held for both temporal and spatial variation. This finding raises the possibility that niche-constructing activity may regulate environmental variables in broadly consistent ways, across diverse traits, taxa, and locations and through time. We suspect that this finding might be a consequence of the weaker selection observed in the constructed category, and we tested this hypothesis through simulation (see figs. S14, S15). The analysis confirmed that the lower between-subset variation observed for constructed and mixed cases could not be explained as a corollary of the substantially reduced within-subset variance, as both kinds of variances are, in principle, independent of each other. However, lowering the average strength of selection directly resulted in reduced variation between subsets. Given that selection gradients in the constructed category are small and often fluctuate around zero, there is little room for differences between different traits and study systems, resulting in reduced between-subset variance.

The finding that niche construction was associated with reduced variation between subsets may also help explain why we did not find strong differences between conditions in our consistency measure. This measure (see eq. [3]) represents the relative proportion of between-subset variation to the sum of between- and within-subset variation. Had niche construction solely reduced within-subset variation, as anticipated, it would have generated greater consistency, but by also reducing between-subset variation, its effect on this measure is less clear.

A potential weakness of the study concerns the accuracy with which sources of selection can be identified by authors. Our methodology does not require all environmental factors to be completely pinned down but rather the general category to be identified, and at least sometimes that is straightforward in practice. Nonetheless, these data will be subject to human error. Ideally, future studies of selection in the wild would record relevant ecological variables, including environmental factors influenced by the activities of organisms, and formally relate these to variation in selection (Endler 1986; Wade and Kalisz 1990; Bassar et al. 2010; MacColl 2011; Caruso et al. 2017). Given our enforced reliance on more circumstantial data to categorize studies and the low number of studies that fulfilled our criteria, our analysis must be regarded as provisional rather than definitive. Until a richer ecologically informed data set can be generated, there will inevitably be concerns over the reliability with which cases can be categorized. These concerns are, to a large part, alleviated by the high interobserver agreement among the two coders conducting this analysis, which demonstrates that by using a well-specified categorization procedure, two individuals will code traits in published studies as constructed and nonconstructed in a consistent fashion. It is also reassuring that simulations found miscategorizing 10% of the data did not greatly affect the conclusions. While such steps guard against bias, they do not preclude error in authors' assessments of the principal sources of selection acting on their study system, which means that our categorization will inevitably be noisy.

Over and above the direct measurement of ecological sources of selection, there are other practical steps that researchers could take to assess the evolutionary impact of niche construction more directly. We were reliant on studies designed to answer different scientific questions from the ones addressed here. While such analyses are clearly instructive, they are no substitute for dedicated experimentation designed directly to test the predictions. An illustration of the kind of experimental investigation required is provided by Nattero et al. (2010), who distinguish between mechanical-fit traits and other flower traits in a specialized pollinator-plant system, reporting weaker selection on the former. Such experimentation clearly separates constructed and nonconstructed cases in the same study, leading to a powerful comparison. A truly satisfactory analysis would require multiple such experiments. Nonetheless, we believe that our meta-analysis is of value precisely because it highlights that there are interesting issues here and encourages the experimentation necessary to resolve them.

Published studies deploying other methods also report findings that raise the possibility of differential selective responses to different sources of selection. For instance, recent studies have demonstrated strong phenotypic changes in organisms in response to urban and other anthropogenic (a case of organism-derived and, hence, constructed but not necessarily regulated) environments compared to natural environments (Palumbi 2001; Palkovacs et al. 2012; Alberti 2015; Alberti et al. 2017; Johnson and Munchi-South 2017; Sullivan et al. 2017). Rather than directly measuring selection, these studies measure phenotypic change, which reflects the effects of selection and phenotypic plasticity and hinders their use to test the present predictions directly. Nonetheless, some of these authors (e.g., Alberti 2015) emphasize the reciprocal relationship between anthropogenic environmental change and distinctive patterns of selection. There is also evidence that in some cases, human-mediated influences can decrease selection by increasing mean fitness (Fugère and Hendry 2018). Such reports further underline the importance of understanding the nature of the source of selection and the possibility that the activities of organisms may shape selection in a manner not yet fully appreciated.

Acknowledgments

We are grateful to Michael Morrissey for assistance with the mathematical models; to Kevin Healy for helpful suggestions concerning the analysis; and to Catherine Sheard, Helen Spence-Jones, Erik Svensson, and Tobias Uller for valuable comments on the manuscript. Research was supported in part by a grant

from the John Templeton Foundation to K.L., J.O.-S., and J.E. ("Putting the Extended Evolutionary Synthesis to the Test"; grant 60501).

Literature Cited

Alberti, M. 2015. Eco-evolutionary dynamics in an urbanizing planet. Trends in Ecology and Evolution 30:114–126.

Alberti, M., C. Correa, J. M. Marzluff, A. P. Hendry, E. P. Palkovacs, K. M. Gotanda, V. M. Hunt, T. M. Apgar, and Y. Zhou. 2017. Global urban signatures of phenotypic change in animal and plant populations. Proceedings of the National Academy of Sciences of the USA 114:8951–8956.

Altman, D. G. 1991. Practical statistics for medical research. Chapman & Hall, London.

Armsworth, P. R., and J. E. Roughgarden. 2005. The impact of directed versus random movement on population dynamics and biodiversity patterns. American Naturalist 165:449–465.

Bassar, R. D., M. C. Marshall, A. López-Sepulcre, E. Zandonà, S. K. Auer, J. Travis, C. M. Pringle, et al. 2010. Local adaptation in Trinidadian guppies alters ecosystem processes. Proceedings of the National Academy of Sciences of the USA 107:3616–3621.

Bogert, C. M. 1949. Thermoregulation in reptiles, a factor in evolution. Society for the Study of Evolution 3:195–211.

Carpenter, B., A. Gelman, M. D. Hoffman, D. Lee, B. Goodrich, M. Betancourt, M. Brubaker, J. Guo, P. Li, and A. Riddell. 2017. Stan: a probabilistic programming language. Journal of Statistical Software 76. https://doi.org/10.18637/jss.v076.i01.

Caruso, C. M., R. A. Martin, N. Sletvold, M. B. Morrissey, M. J. Wade, K. E. Augustine, S. M. Carlson, A. D. C. MacColl, A. M. Siepielski, and J. G. Kingsolver. 2017. What are the environmental determinants of phenotypic selection? a meta-analysis of experimental studies. American Naturalist 190:363–376.

Charmantier, A., L. E. B. Kruuk, J. Blondel, and M. M. Lambrechts. 2004. Testing for microevolution in body size in three blue tit populations. Journal of Evolutionary Biology 17:732–743.

Clark, A. D., D. Deffner, K. Laland, J. Odling-Smee, and J. Endler. 2019. Data from: Niche construction affects the variability and strength of natural selection. American Naturalist, Dryad Digital Repository, https://doi.org/10.5061/dryad.g66n3h5.

Cohen, J. 1960. A coefficient of agreement for nominal scales. Educational and Psychological Measurement 20:37–46.

Combrink, L., H. J. Combrink, A. J. Botha, and C. T. Downs. 2017. Nest temperature fluctuations in a cavity nester, the southern ground hornbill. Journal of Thermal Biology 66:21–26.

Cooper, S. J. 1999. The thermal and energetic significance of cavity roosting in mountain chickadees and juniper titmice. Condor 101:863–866.

Dawkins, R. 2004. Extended phenotype—but not too extended: a reply to Laland, Turner and Jablonka. Biology and Philosophy 19:377–396.

Edelaar, P., and D. I. Bolnick. 2019. Appreciating the multiple processes increasing individual or population fitness. Trends in Ecology and Evolution 34:435–446.

Edelaar, P., A. M. Siepielski, and J. Clobert. 2008. Matching habitat choice causes directed gene flow: a neglected dimension in evolution and ecology. Evolution 62:2462–2472.

Endler, J. A. 1986. Natural selection in the wild. Princeton University Press, Princeton, NJ.

Fugère, V., and A. P. Hendry. 2018. Human influences on the strength of phenotypic selection. Proceedings of the National Academy of Sciences of the USA 115:10070–10075.

Gelman, A., and D. B. Rubin. 1992. Inference from iterative simulation using multiple sequences. Statistical Science 7:457–472.

Hadfield, J. D. 2010. MCMC methods for multi-response generalized linear mixed models: the MCMCglmm R package. Journal of Statistical Software 33:1–22.

Hansell, M. 1984. Animal architecture and building behaviour. Longman, London.

———. 2000. Bird nests and construction behaviour. Cambridge University Press, Cambridge.

———. 2005. Animal architecture. Oxford University Press, Oxford.

———. 2007. Built by animals: the natural history of animal architecture. Oxford University Press, Oxford.

Hawks, J., E. T. Wang, G. M. Cochran, H. C. Harpending, and R. K. Moyzis. 2007. Recent acceleration of human adaptive evolution. Proceedings of the National Academy of Sciences of the USA 104:20753–20758.

Hendry, A. P. 2016. Eco-evolutionary dynamics. Princeton University Press, Princeton, NJ.

Hoekstra, H. E., J. M. Hoekstra, D. Berrigan, S. N. Vignieri, A. Hoang, C. E. Hill, P. Beerli, and J. G. Kingsolver. 2001. Strength and tempo of directional selection in the wild. Proceedings of the National Academy of Sciences of the USA 98:9157–9160.

Huey, R. B., P. E. Hertz, and B. Sinervo. 2003. Behavioral drive versus behavioral inertia in evolution: a null model approach. American Naturalist 161:357–366.

Ieno, E. N., and A. F. Zuur. 2015. A beginner's guide to data exploration and visualization with R. Highland Statistics, Newburgh, NY.

Johnson, M. T. J., and J. Munchi-South. 2017. Evolution of life in urban environments. Science 358:eaam8327

Jones, C. G., J. H. Lawton, and M. Shachak. 1994. Organisms as ecosystem engineers. Pages 130–147 *in* F. B. Samson and F. L. Knopf, eds. Ecosystem management. Springer, New York.

———. 1997. Positive and negative effects of organisms as physical ecosystem engineers. Ecology 78:1946–1957.

Kingsolver, J. G., S. E. Diamond, A. M. Siepielski, and S. M. Carlson. 2012. Synthetic analyses of phenotypic selection in natural populations: lessons, limitations and future directions. Evolutionary Ecology 26:1101–1118.

Kingsolver, J. G., H. E. Hoekstra, J. M. Hoekstra, D. Berrigan, S. N. Vignieri, C. E. Hill, A. Hoang, P. Gilbert, and P. Beerli. 2001. The strength of phenotypic selection in natural populations. American Naturalist 157:245–261.

Laland, K. N., J. Odling-Smee, and J. Endler. 2017. Niche construction, sources of selection and trait coevolution. Interface Focus 7:20160147.

Lande, R., and S. J. Arnold. 1983. The measurement of selection on correlated characters. Evolution 37:1210–1226.

Landis, J. R., and G. G. Koch. 1977. The measurement of observer agreement for categorical data. Biometrics 33:159–174.

Lewontin, R. C. 1983. Gene, organism, and environment. Pages 273–285 *in* D. S. Bendall, ed. Evolution from molecules to men. Cambridge University Press, Cambridge.

———. 2000. The triple helix: gene, organism, and environment. Harvard University Press, Cambridge, MA.

Linnen, C. R., and H. E. Hoekstra. 2009. Measuring natural selection on genotypes and phenotypes in the wild. Cold Spring Harbor Symposia on Quantitative Biology 74:155–168.

Maad, J. 2000. Phenotypic selection in hawkmoth-pollinated *Platanthera bifolia*: targets and fitness surfaces. Evolution 54:112–123.

MacColl, A. D. 2011. The ecological causes of evolution. Trends in Ecology and Evolution 26:514–522.

Morrissey, M. B. 2016. Meta-analysis of magnitudes, differences and variation in evolutionary parameters. Journal of Evolutionary Biology 29:1882–1904.

Morrissey, M. B., and J. D. Hadfield. 2012. Directional selection in temporally replicated studies is remarkably consistent. Evolution 66:435–442.

Morrissey, M. B., L. E. B. Kruuk, and A. J. Wilson. 2010. The danger of applying the breeder's equation in observational studies of natural populations. Journal of Evolutionary Biology 23:2277–2288.

Nattero, J., A. A. Cocucci, and R. Medel. 2010. Pollinator-mediated selection in a specialized pollination system: matches and mismatches across populations Journal of Evolutionary Biology 23:1957–1968.

Odling-Smee, J., D. H. Erwin, E. P. Palkovacs, M. W. Feldman, and K. N. Laland. 2013. Niche construction theory: a practical guide for ecologists. Quarterly Review of Biology 88:3–28.

Odling-Smee, F. J., K. N. Laland, and M. W. Feldman. 1996. Niche construction. American Naturalist 147:641–648.

———. 2003. Niche construction: the neglected process in evolution. Princeton University Press, Princeton, NJ.

Palkovacs, E. P., M. T. Kinnison, C. Correa, C. M. Dalton, and A. P. Hendry. 2012. Fates beyond traits: ecological consequences of human-induced trait change. Evolutionary Applications 5:183–191.

Palumbi, S. R. 2001. Humans as the world's greatest evolutionary force. Science 293:1786–1790.

Pelletier, F., D. Garant, and A. P. Hendry. 2009. Eco-evolutionary dynamics. Philosophical Transactions of the Royal Society B 364:1483–1489.

Pimentel, D. 1961. Population regulation by the genetic feed-back mechanism. American Naturalist 95:65–79.

Plummer, M., N. Best, K. Cowles, and K. Vines. 2006. CODA: convergence diagnosis and output analysis for MCMC. R News 6:7–11.

Post, D. M., and E. P. Palkovacs. 2009. Eco-evolutionary feedbacks in community and ecosystem ecology: interactions between the ecological theatre and the evolutionary play. Philosophical Transactions of the Royal Society B 364:1629–1640.

Ravigné, V., U. Dieckmann, and I. Olivieri. 2009. Live where you thrive: joint evolution of habitat choice and local adaptation facilitates specialization and promotes diversity. American Naturalist 174:E141–E169.

Reid, J. M., W. Cresswell, S. Holt, R. J. Mellanby, D. P. Whitfield, and G. D. Ruxton. 2002. Nest scrape design and clutch heat loss in pectorial sandpipers (*Calidris melanotos*). Functional Ecology 16:305–312.

Scott-Phillips, T. C., K. N. Laland, D. M. Shuker, T. E. Dickins, and S. A. West. 2014. The niche construction perspective: a critical appraisal. Evolution 68:1231–1243.

Siepielski, A. M., J. D. DiBattista, and S. M. Carlson. 2009. It's about time: the temporal dynamics of phenotypic selection in the wild. Ecology Letters 12:1261–1276.

Siepielski, A. M., K. M. Gotanda, M. B. Morrissey, S. E. Diamond, J. D. DiBattista, and S. M. Carlson. 2013. The spatial patterns of directional phenotypic selection. Ecology Letters 16:1382–1392.

Siepielski, A. M., M. B. Morrissey, M. Buoro, S. M. Carlson, C. M. Caruso, S. M. Clegg, T. Coulson, et al. 2017. Precipitation drives global variation in natural selection. Science 355:959–962.

Stan Development Team. 2018. Stan modeling language users guide and reference manual version 2.18.0. http://mc-stan.org.

Stellatelli, O. A., C. Block, A. Villalba, L. E. Vega, J. E. Dajil, and F. B. Cruz. 2018. Behavioral compensation buffers body temperatures of two *Liolaemus* lizards under contrasting environments from the temperate Pampas: a Bogert effect? Ethology Ecology and Evolution 30:297–318.

Sullivan, A. P., D. W. Bird, and G. H. Perry. 2017. Human behaviour as a long-term ecological driver of non-human evolution. Nature Ecology and Evolution 1:0065.

Sultan, S. E. 2015. Organism and environment: ecological development, niche construction, and adaptation. Oxford University Press, New York.

Van Dijk, R. E., J. C. Kadon, A. Arguelles-Toico, L. M. Beltran, M. Paquet, R. Covas, C. Doutreland, and B. J. Hatchwell. 2013. The thermoregulatory benefits of the communal nest of sociable weavers *Philetairus socius* are spatially structured within nests. Journal of Avian Biology 44:102–110.

Wade, M. J., and S. Kalisz. 1990. The causes of natural selection. Evolution 44:1947–1955.

Wadgymar, S. M., D. B. Lowry, B. A. Gould, C. N. Byron, R. M. Mactavish, and J. T. Anderson. 2017. Identifying targets and agents of selection: innovative methods to evaluate the processes that contribute to local adaptation. Methods in Ecology and Evolution 8:738–749.

References Cited Only in the Online Enhancements

Cresswell, W., and R. McCleery. 2003. How great tits maintain synchronization of their hatch date with food supply in response to long-term variability in temperature. Journal of Animal Ecology 72:356–366.

Lill, J. T., and R. J. Marquis. 2003. Ecosystem engineering by caterpillars increases insect herbivore diversity on white oak. Ecology 84:682–690.

———. 2007. Microhabitat manipulation: ecosystem engineering by shelter-building insects. Pages 107–138 *in* K. Cuddington, J. E. Byers, W. G. Wilson, and A. Hastings, eds. Ecosystem engineers: plants to protists. Elsevier, Amsterdam.

Marquis, R. J., and J. T. Lill. 2010. Impact of plant architecture versus leaf quality on attack by leaf-tying caterpillars on five oak species. Oecologia 163:203–213.

Associate Editor: Stephen B. Heard
Editor: Daniel I. Bolnick